The American STATISTICIAN

The news publication of the
AMERICAN STATISTICAL ASSOCIATION

MAR 1 '50

NEW BASE PERIOD FOR ECONOMIC INDEXES
A letter to the editor

NEWS 2, 5

OVERLAPPING MEMBERSHIP OF SOCIETIES
Dickson H. Leavens 4

THE DEVELOPMENT OF FEDERAL STATISTICAL COORDINATION — Part II Clem C. Linnenberg, Jr. 6

THE SUPPLEMENTARY SCALE CHART Kenneth W. Haemer 11

QUESTIONS AND ANSWERS The Proposed Question
edited by Frederick Mosteller 12

CHAPTER NOTES 16

JUNE-JULY 1949
Volume 3, No. 3

25 CENTS

MONTANA STATE COLLEGE LIBRARY
BOZEMAN

Statistics in relation to . . . communication engineering

THE EXTRAPOLATION, INTERPOLATION, and SMOOTHING of

STATIONARY TIME SERIES

with ENGINEERING APPLICATIONS

By **NORBERT WIENER**, Professor of Mathematics, The Massachusetts Institute of Technology

Time Series represents an attempt to unite the theory and practice of two vitally important fields of work—time series in statistics and communication engineering. The timeliness and importance of this book is evidenced by the fact that in an increasing number of organizations (as the Bell Telephone Laboratories and M.I.T.) there is a tendency toward the revision of the whole communication engineering field as a branch of statistical mechanics.

In all communication engineering the message to be transmitted is represented as some sort of array of measurable quantities distributed in time. In other words, by coding or the use of the voice or scanning, the message to be transmitted is developed into a time series. It is such a relationship that is discussed in this new book. *Time Series* introduces methods leading from existing techniques of both the statistician and engineer and fuses them into a common technique which is more effective than either one alone.

A TECHNOLOGY PRESS PUBLICATION

August 1949 163 pages $4.00

. . . and psychology

PSYCHOLOGICAL STATISTICS

By **QUINN McNEMAR**, Professor of Psychology, Statistics, and Education, Stanford University

Psychological Statistics answers the need for a concise introduction to statistical methods with a continuous transition to more advanced topics. Except for factor analysis, all the techniques frequently used in psychological research are covered.

Psychological Statistics begins with a discussion of elementary material and proceeds to an extensive exposition of sampling and of correlational analysis. Then follows a thorough treatment of that part of statistical inference which involves chi square, small sample methods, and the analysis of variance. An understanding of fundamentals is particularly stressed, with emphasis on assumptions, permissible interpretations and inferences, and limitations.

A Wiley Publication in Psychology; Herbert S. Langfeld, Advisory Editor

January 1949 364 pages $4.50

Copies available on approval

JOHN WILEY & SONS, Inc., 440–4th Ave., New York 16, N. Y.

The American STATISTICIAN

JUNE-JULY 1949, VOL. III, NO. 3

The news publication of the
American Statistical Association

OFFICERS

President: Simon Kuznets. *Directors:* C. H. Goulden, Isador Lubin, Lowell J. Reed, George W. Snedecor, Frederick F. Stephan, Willard L. Thorp, L. L. Thurstone, Samuel S. Wilks. *Vice-Presidents:* Dorothy S. Brady, H. A. Freeman, Lester S. Kellogg. *Secretary-Treasurer:* Samuel Weiss.

EDITORIAL COMMITTEE

Sylvia Castleton Weyl, editor; Churchill Eisenhart, Morris Hansen, Samuel P. Hayes, Jr., Lester S. Kellogg, Julius Lieblein, Will Lissner, Marion M. Sandomire, John W. Tukey, Morris Ullman.

Department Editors:

Questions and Answers
Frederick Mosteller

Presentation Problems
Kenneth W. Haemer

Correspondents: Albany — David M. Schneider; Austin — Stella Traweek; Central New Jersey—Willard C. Thompson; Chicago—Jack R. Green; Cincinnati—Ellery F. Reed; Columbus—Rosemary Tague; Connecticut—J. H. Watkins; Cuba—Hugo Vivo; Philadelphia—Mary McDermott; New York—Milton Stevens; Washington, D. C.—Marguerite Gurney; Government—Virginia T. Venneman; Latin America—Francisco de Abrisqueta; United Nations—F. Marguerite Nowak.

Entered as second class matter March 11, 1938, at the post office at Washington, D. C., under act of March 3, 1897. *The American Statistician is published six times a year*—February, April, June, August, October and December—by the American Statistical Association, Editorial Office: 1603 K Street, N. W., Washington 6, D. C. Subscription rate: one dollar and fifty cents a year or twenty-five cents per copy.

LETTER TO THE EDITOR

New Base Period for U. S. and Regional Economic Indexes

It is obvious to all who are working with index numbers of business and economic data that the adoption of a new base period is eminently necessary. The pre-war, 1935-1939 widely-used base period is already more than a decade in the past, and represents a period in which the price levels and business activity are hardly representative of current economic levels. The current business and politico-economic framework is vastly different from that of the middle '30's and hence, it seems ill-advised to maintain the latter period as a more or less point of reference. Present levels are at 300, 400, and 500 magnitudes in many value series, which in itself is at least a partial contradiction of the index idea of ratios quickly comparable with the 100 base.

As far as we have been able to determine, there is at present no Committee of the American Statistical Association canvassing the problem of new index base periods in order to make official recommendations.

If such a Committee is not already functioning, we suggest that a Committee be appointed for this purpose. We further suggest that it should be the duty of the Committee to: (1) confer with the chiefs of the various statistical agencies of the U. S. Government and with private research organizations, trade associations and state agencies which compile and publish index numbers, in order to arrive at a consensus on the best new base period to be adopted for all U. S. business and economic index numbers; and (2) that the Committee function through the American Statistical Association as an educational agency to secure the universal adoption of the recommended base period by all agencies compiling and publishing index numbers. The committee, we believe, should be composed of representatives from professional business statisticians in the universities, U. S. Government agencies, research foundations and state agencies.

Statisticians understand that the base period of an economic index is simply an arbitrarily chosen point of reference and has nothing to do with "normal" in an economic sense, nor with a social norm. Practically, however, it must be recognized that whatever period is selected will be popularly conceived of in terms of "normality". The likelihood of such an interpretation of the base period should be a consideration in its choice.

For our own part we should like to recommend strongly to the consideration of such a Committee that a *single year* be adopted as the universal base period to be recommended by the Committee. The advantages of a single year in ease of calculation are so compelling as not to be overlooked, particularly in view of the enormous job of converting to the new base.

We should like to suggest further, that as *the* single year, 1950 is a "natural." Barring any possible collapse before that time, 1950 will likely represent a year in which economic conditions will have "settled down." Also, 1950 will about coincide with the availability of data from censuses now in progress or which will be taken as of that year. Thus, adjustment of indexes to census bench marks and conversion to the new base period can be made simultaneously. Finally, it is the mid-century year, for whatever value that may have.

JAMES C. YOCUM
SAM ARNOLD
Bureau of Business Research
The Ohio State University

NEWS

Nominations for 1950 officers of ASA—Presentation clinic at annual meeting—U. K. statistical organization—Slide rule for the blind—Civil Service examinations—Committee appointments—FAO census—new publications

New committee appointments for 1949

Business Statistics Section

Ralph J. Watkins, *Chairman.*

Henry Arthur, Robert Burgess, Richard Funkhauser, Martin Gainsbrugh, Lester Kellogg, and Benedict Saurino.

ASA Subcommittee on BLS Price Indexes

An American Statistical Association Subcommittee on the Bureau of Labor Statistics Price Indexes has just been established. This group is being appointed as a subcommittee of the ASAA Commission on Statistical Standards to advise on the revision of the BLS price indexes. The members of the ASA subcommittee are: Dr. Reavis Cox, University of Pennsylvania; Dr. Dudley J. Cowden, University of North Carolina; and Dr. Bruce D. Mudgett, University of Minnesota.

1949 Annual Meeting Program

Helen Slade has accepted the position of ASA chairman for local arrangements for the 1949 Annual Meeting to be held in New York December 27 through 30, 1949.

Presentation Clinic at Annual Meeting

As a special feature of the Annual Meeting, a presentation clinic has been organized by the ASA Committee on Presentation.

Under the chairmanship of John W. Tukey, Princeton University, the panel will consist of the following clinicians: Rudolf Flesch, Author and Consultant; Kenneth W. Haemer, American Telephone and Telegraph Company; Bruce L. Jenkinson, Bureau of the Census; Henry Margenau, Yale University; Rudolf Modley, Executive Research, Inc.; William A. Neiswanger, University of Illinois; Walter A. Shewhart, Bell Telephone Laboratories; W. J. Youden, National Bureau of Standards.

This Panel represents a broad cross-section of experience in presentation problems and their solution, and it is expected that a wide variety of presentation topics will be discussed at this session. Although it may be possible to accept some questions from the floor, most of the time will be devoted to answering questions submitted in advance.

Questions should be addressed to the Secretary of the American Statistical Association, 1603 K Street, N.W., Washington 6, D. C., and should be sent to arrive not later than November 15th.

Statistics

Published in February, a new Occupational Abstract titled "Statistics" is now available from Occupational Index, Inc., New York University, New York 3, New York, for 50 cents a copy, cash with order. This abstract was written by Josephine McGinness and contains occupational information of value to the high school counselor and student.

Included in this 6-page pamphlet are a description of the work at various levels, qualifications, preparation, future prospects, unions and discrimination, methods of entrance and advancement, earnings, number and distribution of workers, women in the field, advantages and disadvantages. There is also information on the available literature, sources of further information, and a brief bibliography.

A slide rule for the blind

A new braille slide rule has been devised recently by the T. R. Dolftee American Foundation for the Blind, Inc. The slide rule is circular, with a logarithmic "C" scale engraved around the outer edge and braille indications at each tenth placed on the back. The rule has three plastic pointers, two, A and B, on the face, and a third, C, on the back, rigidly attached to A. Pointer A carries pointer B along with it, but pointer B may be moved without shifting pointer A. The use of pointer C, on the back, enables the operator to take readings on a linear scale.

The braille circular slide rule would appear to be well designed for the operations of multiplication, division, and taking roots, by sightless engineers and students. It should extend the advantages of rapid low-precision computation to a class of very earnest students.

U. K. Central Statistical Organization

The need for a Central Statistical Organization in Great Britain has been advocated for some 70 years. Its actual establishment in 1941 was a result of the war and of the increasing wartime emphasis on statistics as a technique of administrative control, production performance, and planning.

Central statistical organizations were set up in many departments and given responsibility for the collection of data. The Central Statistical Organization was placed in the Offices of the War Cabinet and from the outset served as a coordinating agency. Decisions had to be made continuously which affected the work of several cabinet departments. These decisions were predicated on statistical data collected by independent government agencies. One of the principal jobs of the CSO was to ensure that the data were compiled in a consistent manner and on the basis of agreed-upon definitions.

The main functions of the Central Statistical Organization at present are:

a) To assist the Cabinet Secretariat on all statistical questions and in the preparation of statistical reports and collection of information required by the Cabinet;

b) To maintain general liaison between the Cabinet Secretariat and the individual departments;

c) To centralize preparation and circulation of all reports containing data obtained from more than one department and to assist in the preparation of such reports;

d) To advise the departments on statistical matters and to arrange interdepartmental discussions on statistical problems.

Contact between the CSO and the several departments is maintained on a daily and ad hoc basis. Informal interdepartmental committees have been set up on such continuing problems as Balance of Payments, National Income, Trade Accounts, U. N. Relations, Employment Policy, and Industrial Classification.

The CSO is a skeleton organization (by American standards) with 25 administrative and executive officers and 30 clerical. Each administrative officer is responsible for day-to-day liaison with the statistical divisions of the departments and keeping informed on all policy matters within his sphere of responsibility.

During the war and postwar periods, the use of economic and statistical material in Government increased enormously. A more ambitious program of statistical collection and compilation became necessary. Accordingly, the 1947 Statistics of Trade Act provides, interalia, for an annual census of production and a periodic census of distribution.

Among the regular publications of the CSO are: *Monthly Digest of Statistics*, the *Annual Abstract of Statistics* and the annual *White Paper on National Income* (published each April). The above are obtainable from the British Information Services in New York.

New quarterly in India official publication of the Indian Society of Agricultural Statistics

A new publication, "Journal of the Indian Society of Agritultural Statistics" has just been received. This journal is the organ of the Indian Society of Agricultural Statistics and each volume will be issued in two parts annually.

The first issue contains the presidential address at the first meeting of the Society, a symposium on "Statistical Organization for India with Special Reference to Agriculture;" "Crop Surveys in India—I" by V. G. Panse and P. V. Sukhatme; "A New Approach to Sampling Distributions of the Multivariate Normal Theory—I" by R. D. Narain; "On the Distribution of Estimated Error Components in the Analysis of Variance and Covariance" by R. D. Narain; "Estimation of Genetic Variability in Plants" by V. G. Panse and S. D. Bokil; "On Fractional Replication of the General Symmetrical Factorial Design" by K. Kishen.

Books and periodicals for review should be sent to the Editor, Dr. V. G. Panse, Institute of Plant Industry, Indore.

Report of the Committee on Elections

In accordance with the provisions of Section 3, Article VII, of the Constitution, the Committee on Elections of the American Statistical Association submits the following slate of nominees:

President-Elect	Lowell J. Reed
Vice-President (3 years) (one to be elected)	Robert W. Burgess Philip M. Hauser
Directors (3 year term) (two to be eletced)	Gertrude M. Cox W. Edwards Deming Frank R. Garfield Holbrook Working

The Constitution specifies that:

"These nominations shall be published in the next news bulletin. Additional nominations may be made within five weeks after this publication by petition signed by at least 25 members and submitted to the Secretary-Treasurer.

"On or before November 15, the Secretary shall mail to the members of the Association a brief biographical sketch of each nominee with a preferential ballot for the election of officers from among the persons so nominated."

Respectfully submitted,

JOSEPH S. DAVIS
ALFRED N. WATSON
SAMUEL S. WILKS, *Chairman*

Nominations for District Representative to the Council come directly from the membership

Article VII, Point 3 of the Constitution states "For district representative, any member may submit a nomination for his district to the office of the Secretary-Treasurer in writing . . . The four nominees receiving the largest number of nominations in their respective districts, as determined by the Committee on Elections, shall be the nominees for their district".

Civil Service examinations

Recent information obtained from the Civil Service Commission indicates that another Junior Professional Assistant examination probably will be given early this autumn, possibly in October. This examination will provide an option for Junior Statisticians. Another examination which probably will be given at about the same time is for Junior Agricultural Assistant with an option for Junior Agricultural Statistician. These forthcoming examinations should be of particular interest to students. Under Civil Service rules, students who are in their Senior year and who expect to complete work for a Bachelor's degree before June 30, 1950, are eligible to file for these examinations. Members of the Association engaged in the teaching of statistics should urge promising students to file for these examinations. As soon as definite information regarding the forthcoming JPA examinations or other examinations is available, it will be reported.

The Civil Service Commission has indicated that at the present time there are no examinations open for statisticians, except in the 4th Civil Service Region which has one examination open for statisticians, Grades P-2 through P-5 for certain military establishments located near Washington, D. C.

CONTINUED ON PAGE 5

Overlapping Membership of Societies

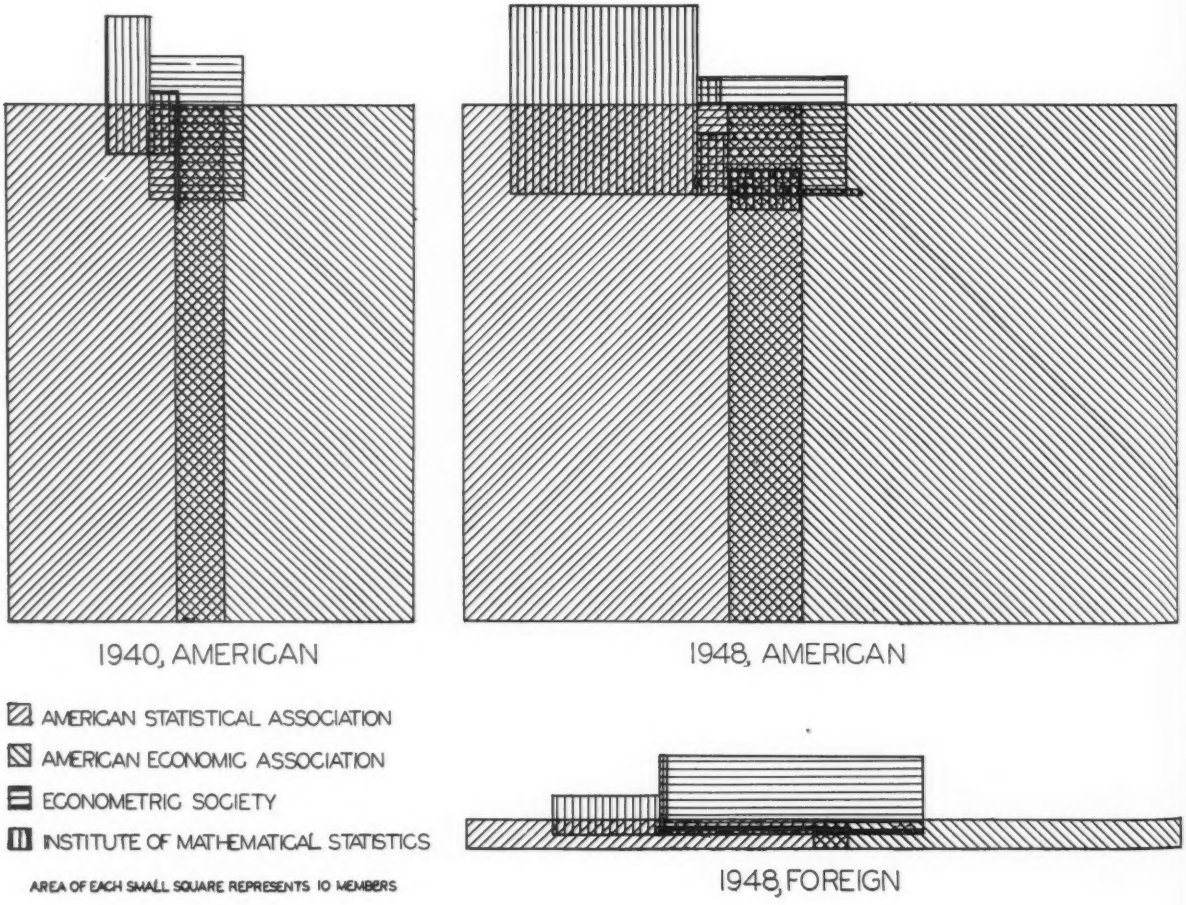
by DICKSON H. LEAVENS
Colorado Springs, Colorado

In 1941 the writer cross-checked the membership lists, published at the end of 1940 or early in July, of the American Statistical Association, the American Economic Association, the Econometric Society, and the Institute of Mathematical Statistics, and summarized the number of overlapping memberships, giving copies to the secretaries of the various societies.

A similar check has now been made for 1948 memberships. These were available in published form for two of the societies: *American Economic Review*, Vol. 39, No. 1, January, 1949, giving membership as of June 15, 1948; and *Econometrica*, Vol. 16, No. 4, October, 1948, pp. 377-402, giving membership as of October 31, 1948. The other two societies kindly supplied addressograph lists early in January, 1949.

Because of the amount of time required, no attempt has been made to recheck the work, and the figures in some categories may be a few units off, but they should give an essentially correct picture of the overlapping membership.

The results are summarized in the accompanying table and shown graphically in the chart. For the oldets three societies it is easy to show the exact status by three overlapping rectangles, but when a fourth is added a rectangle cannot show all the results and it is necessary to use an irregular figure. The 1940 survey



included only members in the United States, because, except for the international Econometric Society, foreign members made up only a small part of the total. For 1948 the foreign members have been shown separately. The criterion for "American" and "foreign" has been the address shown, but it is probable that Americans temporarily abroad are about balanced by foreigners temporarily in this country.

As might be expected, the Statistical and Economic Associations each has a large majority of American members who do not belong to any of the other associations; while the Econometric Society and the Institute, being more specialized, have a majority of American members (80 per cent for E. S. and 56 per cent for I. M. S.) who belong to one or more other societies.

It may be of interest to note the change from 1940 to 1948 in the percentage of membership of the American Economic Association who belong to some of the three organizations that are devoted to quantitative methods. In 1940, 27 per cent belonged to one or more of the other three organizations; in 1948 this had increased to 33 per cent. This would seem to indicate a growing interest among economists in quantitative methods. On the other hand the percentage of members of the American Statistical Association who belong to either or both the economic groups has decreased from 24 per cent to 22.5 per cent, reflecting the increasing use of statistics in other fields.

Among the foreign members, there is considerable overlapping, but the Econometric Society, being an international organization with more foreign than American members, has a large group who have not been drawn into any of the three American societies, although most of them probably belong to economic and statistical organizations in their own countries.

Overlapping Memberships

	1940 American	1948 American	1948 Foreign
Total memberships:			
American Statistical Association	2625	4087	250
American Economic Association	2889	5484	245
Econometric Society	303	416	466
Institute of Mathematical Statistics	182	964	143
Grand total memberships	5999	10951	1104
Distribution of memberships:			
Members of one society:			
Statistical alone	1928	2736	154
Economic alone	2251	4493	202
Econometric alone	88	82	378
Institute alone	88	439	65
	4355	7750	799
Members of two societies:			
Statistical, Economic	483	714	11
Statistical, Econometric	26	22	19
Statistical, Institute	49	390	32
Economic, Econometric	42	89	18
Economic, Institute	.....	2	.....
Econometric, Institute	9	17	12
	609	1234	92
Members of three societies:			
Statistical, Economic, Econometric	103	115	5
Statistical, Economic, Institute	1	25	.....
Statistical, Econometric, Institute	26	45	25
Economic, Econometric, Institute	.....	6	5
	130	191	35
Members of all four societies	9	40	4
Total different persons in memberships	5103	9215	930

NEWS CONTINUED FROM PAGE 3

Norway and Bizone Germany launch FAO's 1950 World Census of Agriculture

Norway is the first member government to launch FAO's 1950 World Census of Agriculture. Census workers of that country began on Monday, June 20, to gather a mass of information which will give an accurate picture of Norway's agriculture, forestry, and fisheries.

The information thus gained will be used by FAO as part of a statistical report on world agriculture more comprehensive than has ever been attempted before.

The world census of agriculture is not confined to FAO's 58 member countries, however. It is hoped that all countries will participate. FAO officials are pleased to note in that connection that the Bizone of Germany was actually the first area to take the census. On May 22 80,000 census workers set out to interview all farmers cultivating more than half a hectare of land. There are about 1,600,000 such farmers in the area. Exact information as to the material included in the census list has not yet been received, but it is understood that a comprehensive list of questions was asked.

A broad list of questions is included in the Norway census. Among items on which information will be assembled are area in crops in 1949; other land uses, such as meadows for pasture, forests, and fallow land; areas in vegetables; area under glass; the number of livestock; the number and types of machines in use on farms; the tenure of farm operators; the amount of time worked off the farm by the operator; whether the farm wife works in the fields; whether the farm operator is a trained agriculturalist; the characteristics of labor on farms—sex, age, wages paid, etc.; farm plumbing and wiring; farm expenses, drainage and irrigation practices, and the use of silage.

Traditionally, Norway has operated an excellent agricultural statistical service. Its present census list is broad enough to include practically all the information which FAO hopes, through the 1950 census, to obtain country by country. The information sought from each country, FAO believes, is of indispensable assistance in national and international planning for increased agricultural production and improved distribution of food and other essential products of agriculture.

For purposes of international comparability, crop data gathered for the Northern Hemisphere, FAO has suggested, should refer as far as possible to crops harvested in the calendar year 1949 and, in the case of some parts of the Far East to harvests continuing into the early part of 1950. For the Southern Hemisphere, data should refer to crops harvested in the year from July 1, 1949, to June 30, 1950.

Most countries, it appears, will wait until 1950 to get the census under way. In addition to Norway and Bizonal Germany, however, Denmark, the French Zone in Germany and Austria plan to conduct the census in the present calendar year.

The Development of Federal Statistical Coordination, 1908-1949

by CLEM C. LINNENBERG, Jr.*
Division of Statistical Standards
U. S. Bureau of the Budget

(Continued from the April issue)

Division of Statistical Standards, 1940 to the present

a. Reliance upon the Budget and Accounting Act.

The Budget and Accounting Act, enacted in 1921, created the Bureau of the Budget. As was indicated in the first installment of this article, the Central Statistical Board was placed in that Bureau in 1939; and the Board's staff, the statistical coordinating group of the Federal Government, had already begun to operate under the Budget and Accounting Act before July 25, 1940, when Public Law 219 and the Central Statistical Board expired. The task of coordinating Federal statistics has also come to rely substantially upon the Budget and Accounting Act in ways not discernible in Circular No. 360 or in any other regulation. The Budget Bureau's Division of Estimates and its Division of Statistical Standards consult each other on innumerable occasions, usually in an informal manner, regarding Federal statistical activities. In the development of the annual Federal budget, each agency submits estimates (requests for funds) to the Budget Bureau. The Bureau's Division of Estimates holds hearings to determine the reasonableness of the sums requested. At such a hearing, the representatives of a given agency appear, make oral statements in support of their budget estimates, and are questioned. The Division of Estimates invites Division of Statistical Standards people to sit in on these hearings, allowing them to participate in the questioning of agency representatives, particularly those who are connected with reports and statistics, and to some extent asking these D. S. S. people to participate in staff discussions thereafter, where the testimony is evaluated. Similarly, the D. S. S. asks Estimates Division examiners to participate fully in any D. S. S. meetings with statistical-agency personnel that are likely to interest the examiners—such as a meeting regarding a statistical proposal which is likely to be expensive. Each division is responsible for its own decisions, but each is helped by the other's experts. The result is a substantial if sometimes subtle influence upon the statistical activities involved.

b. Budget Bureau Approval Numbers.

(1) Under the Central Statistical Board, and under the Bureau of the Budget for more than three years, there was nothing printed on Federal forms to indicate to the public that they had been reviewed by the statistical coordinating agency. (2) In any event, completion of review had never been regarded—by the public, by the data-collecting agency, or by the statistical coordinating body—as involving *approval* by this last-named entity. (3) Furthermore, review had ordinarily occurred only on adoption or revision of a form, with no reconsideration from time to time unless the using agency proposed a revision. The perennial public unhappiness over filling out government forms was markedly increased by the large volume of applications, reports, and the like which the public was required to fill out in connection with wartime economic controls. The Bureau of the Budget decided to remedy these three deficiencies in the scheme of statistical review. It was thought that an indication of approval by the Budget Bureau would be some assurance to the public (particularly to the businessmen who could be informed of this device through their trade journals) that any form thus labeled had been carefully considered by an impartial, central agency, for lucidity of wording, avoidance of duplication and of needless questions, and the like. It was also believed that, to the extent that people would come to look for an indication of Budget Bureau approval on Federal forms, the using agencies would feel impelled to comply with the requirement of review, and to accept enough of the suggestions made during review to ensure Budget Bureau approval for their forms.

On August 31, 1942, the Bureau issued Supplement No. 1 to Circular No. 360. This declared: "Effective immediately, the Bureau of the Budget will begin assigning approval numbers to all report forms used by Federal agencies (except oral or telegraphic requests for information) which have been reviewed and cleared by the Bureau. Federal agencies are hereby authorized to print or otherwise reproduce upon the copies of a report form the approval number assigned to it . . . After January 1, 1943, every report form used by any agency shall bear an approval number assigned to it by the Bureau of the Budget . . ." Meanwhile, all repetitive forms (forms other than for a single-time survey) which were already in use were to be submitted for assignment of approval numbers. For any given form assigned an approval number either in this transition period or subsequently, the number would possibly have "an expiration date after which the approval of the form . . . [would] automatically be withdrawn." For use of the approval number after the expiration date, a new submittal, review, and clearance would be re-

*The views expressed in this article are not necessarily those of the agency.

quired. Submittal anew was also to be required in the event that the report form was to undergo a substantial revision or change in use.

Presumably the amended Circular No. 360's prohibition of the use of forms without Budget Bureau approval numbers was based at least in part upon Executive Order No. 8248, Section II-2(g), quoted in the *American Statistician*, April 1949, p. 10.

It was such a huge task to initiate the approval-number system for the accumulation of forms already in use, that it was indeed fortunate that it was undertaken at this particular juncture. On December 24, 1942, the Federal Reports Act of 1942 became law. This gave to the Budget Bureau wide powers of coordination over the fact-gathering activities of the Federal Government. In the exercise of one particular part of this authority (Sec. 5), the approval-number system is extremely helpful. The sheer mechanics of simultaneously reviewing all repetitive Federal forms already in use and assigning approval numbers to most of them was so time-consuming that it would have greatly slowed down the Bureau's getting under way, on a truly *substantive* level, in carrying out the strenuous duties placed upon it by Congress in that part of the Federal Reports Act. However, by December 24 most of the task of getting the approval-number system into operation had been accomplished.

c. Origin of the Federal Reports Act.

The principles to which Congress gave its sanction in the Federal Reports Act had been emerging slowly for some years. In telling about them, a good starting point is the "Report of the Central Statistical Board on Returns Made by the Public to the Federal Government." (House Document No. 27, 76th Congress, 1st Session.) This study, prepared on special request of President Roosevelt in 1938, presented descriptive and quantitative data about reports and returns, both administrative and non-administrative, which are submitted by the public to the Federal Government. It concluded with "An Immediate Program for Improved Coordination." Most parts of this were proposals of new legislation—not coordination by act of Congress, but measures such as an authorization to the statistical coordinating agency to designate one agency to serve as the data-collecting office for two or more agencies in a specified field which is of interest to all of them.

A half-year after this study appeared, the Central Statistical Board became a part of the Bureau of the Budget, as noted above, and there was thus an assurance that at least the existing degree of statistical coordination would continue. However, the Federal quiz program still had its excesses. The Senate Special Committee to Study Problems of American Small Business, and especially the Committee Chairman, Senator Murray, showed a lively interest in this matter, and a desire to help businessmen without harming any proper Federal activity. It used as its point of departure the "Report . . . on Returns Made by the Public . . ." and the replies which, in a closely related survey, had been received by the White House from the Federal agencies most concerned, in response to questions about statistical coordination. "These replies . . .," said the Committee, "almost unanimously rejected the idea of a consolidation of Federal statistical work into one agency. On the other hand, they just as unanimously advocated the establishing by law or otherwise of an agency to coordinate Federal reporting services and eliminate duplications . . ." In June, 1941, the Committee proposed a Federal Reports Act, conferring wide powers of statistical coordination upon the Budget Director. The bill (S. 1666 and H. R. 5180) was referred at that time to the Senate Committee on Education and Labor and the House Committee on Expenditures in the Executive Departments.

The bill remained dormant for a year, but meanwhile the defense agencies' needs for information were causing, as noted above, an increased demand that something be done to keep within reason the burden which Federal requests for data were producing. In June, 1942, the Senate Committee on Education and Labor set up a subcommittee, headed by Senator Murray, to consider the bill anew. The result of this action and of attention by the House Committee on Expenditures in the Executive Departments, which also considered a companion measure, H. R. 7756, was that a bill conferring even more comprehensive powers of statistical coordination upon the Budget Director became law on Christmas Eve as the Federal Reports Act of 1942 (5 U. S. C. 139-139f).

d. Powers Granted in the Federal Reports Act.

Thus far, this has been the only permanent Federal legislation intended specifically for the coordination of statistics. Moreover, while the measures taken in the preceding decade had given more or less comprehensive *responsibility* to a coordinating agency, this law gave such an agency both *responsibility* and *authority* on a comprehensive scale. The Act has none of the three inadequacies referred to in the discussion of the Central Statistical Board statute, Public Law 219. (See the *American Statistician*, April 1949, p. 9.)

Section 2 of the Federal Reports Act declared a three-fold policy of Congress regarding information: (1) Holding to a minimum the burden upon the public in the furnishing of information to the Federal Government. (2) Holding the cost to the Government to a minimum. (3) Tabulation of information by any Federal agency in a manner to make the data useful to other Federal agencies and the public. Incidentally to (1) and (2), it was declared that all unnecessary duplication of effort in obtaining in-

formation should be eliminated as rapidly as practicable.

The powers conferred by the Act upon the Budget Bureau include (a) an authority which automatically comes into operation, and inevitably figures in every single day's work of the Bureau, and (b) a group of powers, mostly operative only on the Bureau's own motion, which do not need to be exercised so frequently.

The coordinating role given to the Budget Bureau by this law applies to all Federal executive agencies (departments, regulatory commissions, etc.) with a few exceptions, chiefly some bureaus of the Treasury Department (Secs. 3 (e) and 7 (a)). The Act defines "person" very broadly, to include any individual, corporation, organized group, or State, Territorial, or local government (Sec. 7 (b)). It defines "information" as "facts obtained or solicited by the use of written report forms, application forms, schedules, questionnaires, or other similar methods calling either (1) for answers to identical questions from ten or more persons other then [means 'than'] agencies, instrumentalities, or employees of the United States or (2) for answers to questions from agencies, instrumentalities, or employees of the United States which are to be used for statistical compilations of general public interest" (Sec. 7 (c)). It should be observed that part (1) of this definition avoids any reference to information as being statistical or non-statistical, and any reference to the purpose which the information is to serve—administration, research, etc. Avoiding such references is a neat way of averting a great deal of argument over jurisdiction. And it happens that (1) is the part of the definition which concerns getting information from the public, and which is involved in statistical coordinating activities hundreds of times for every occasion when (2) is involved. In my discussion of the Reports Act and of the Budget Bureau's work under it, I shall use the word "information" as defined in the Act.

Sec. 5 embodies the day-to-day sort of authority mentioned above. It provides that no Federal agency covered by the Act shall either conduct or sponsor the collection of "information," as defined by the Act,

> . . . unless, in advance of adoption or revision of any plans or forms to be used in such collection, (a) the agency shall have submitted to the Director [of the Bureau of the Budget] such plans or forms, together with copies of such pertinent regulations and other related materials as the Director shall specify; and (b) the Director shall have stated that he does not disapprove the proposed collection of information.

In this section, Congress made the legality of innumerable executive fact-gathering activities hinge in part upon Budget Bureau assent. As a result, there is a daily flow of information-collecting proposals, of diverse purpose, scope, and merit, into the Budget Bureau, and a daily flow outward of actions upon such proposals. For years before, there had been a daily two-way flow; but the Bureau is in a position now to take definitive action.

Sec. 5 alone would have been a terrific innovation. The other powers are no less impressive. Except as noted below, these are not intended to apply to daily operations. Secs. 3 (a)-3 (c) require the Budget Director, "With a view to carrying out the policy of this Act," to investigate from time to time the needs of the various Federal agencies for "information," and their methods used in obtaining it, and to coordinate their information-collecting services as rapidly as possible. After such an investigation, if he decides that a single collecting agency would adequately serve the needs of any two or more Federal agencies for information, he shall hold a hearing

> . . . . at which the agencies concerned and any other interested persons shall have an opportunity to present their views. After such hearing, the Director may issue an order designating a collecting agency to obtain such information for any two or more of the agencies concerned, and prescribing (with reference to the collection of such information) the duties and functions of the collecting agency so designated and the Federal agencies for which it is to act as agent . . . [Thereafter] no Federal agency covered by such order shall obtain for itself any information which it is the duty of the collecting agency designated by such order to obtain.

The idea behind this piece of authority can be found not only in the special report prepared for the President in 1938 but also in the *First Annual Report of the Central Statistical Board,* where it was urged that in each field there be a "focal agency," to collect information peculiar to that field (and to do research and analysis based on these data and on statistics borrowed from activities, such as the censuses, which cut across fields). (C. S. B. *Report,* for calendar year 1934, pp. 8-10.)

Section 3 (d) of the Federal Reports Act authorizes the Budget Director to determine

> . . . whether or not the collection of any information by any Federal agency is necessary for the proper performance of the functions of such agency or for any other proper purpose.

If he finds that the information is not needed or is obtainable from another Federal agency, the agency doing the unnecessary collecting must cease to do so.

However, it is small comfort to an agency to know that a second agency possesses data if the latter will not disgorge them or is forbidden by law to do so. Secs. 3 (e) and 4 face up to this fact. Sec. 3 (e) authorizes the Budget Director "to require any Federal agency to make available to any other Federal agency any information which it has obtained from any person after the date of enactment of this Act . . ." Sec. 4 recognizes that much information is gathered on a confidential basis, as a protection to trade secrets, personal feelings of the public, etc. E.g., Sec. 4 (b) limits 3 (e) by providing that, in activities pursuant to this law, confidential information obtained

by a Federal agency is not to be released to another Federal agency unless at least one of several alternative conditions is met, such as that the persons originally supplying the information "consent to the release of it to a second agency by the agency to which the information was originally supplied."

Sec. 9 of the Reports Act authorizes the annual appropriation of ". . . such sums as may be necessary to carry out . . ." the Act's provisions. This is, of course, no assurance as to how much money will be provided each year. But it does mean that Congress has left itself free to consider, on the merits, each year's request for funds for the statistical coordinating staff, instead of fixing an exact ceiling on a permanent basis.

e. Procedures Established by Budget Circular No. A-40.

Sec. 6 of the Federal Reports Act is potentially relevant to both sorts of powers which the Act contains—those used every day and those for occasional use. This section authorizes the Budget Director to issue regulations to carry out the Act's provisions. Only one has been issued—"Clearance of plans and report forms under the Federal Reports Act." This was issued in February, 1943, as Regulation A and, slightly revised, was re-issued in October, 1948, as Budget Circular No. A-40. It has considerable similarity to Regulation No. 1 and Budget Circular No. 360. It replaced the latter.

By means of Circular No. A-40, the Budget Director delegated to the Division of Statistical Standards his authority under the Reports Act's Sec. 5 (the type of authority intended for day-to-day use) and under Sec. 3 (d) (see above). So far as concerns the agencies covered by the Act—and this means nearly all of the executive agencies—the circular controls various devices used by the Government in gathering "information," no matter whether the data are provided by the public voluntarily or under compulsion. The devices covered by the circular are forms and anything similar, such as standardized interviews for fact-gathering; regulations calling for the submittal of data; and regulations calling for the maintenance of records from which data are to be submitted later to the Government. No such device is to be used in connection with obtaining "information" unless the responsible agency is authorized in advance by the Division of Statistical Standards. Clearance by the Division is usually for a specified period of time, even if the reporting scheme is intended by the responsible agency to continue permanently. A new authorization by the D. S. S. is needed for use after such an expiration date, or if the form or other device is to be materially revised. In any event, the D. S. S. may withdraw the authorization at any time, despite a previously assigned expiration date or the lack of one.

Circular No. A-40 resembles No. 360 in the breadth of the devices (forms and regulations) which are covered. It resembles both No. 360 and the still earlier Regulation No. 1 as to general purpose, and as to the sort of supporting statement which must accompany each proposal submitted for review. However, the requirement as to a supporting statement now reflects a long experience with statistical coordination. (For instance, if sampling is not to be used, the reason for complete coverage must be supplied.) The type of supporting statement also reflects the fact that that review no longer relies solely upon the coordinating staff's technical competence and its ability at persuasion but upon a Congressional grant of genuine authority as well. Part of the supporting statement is called a "statement of *justification*" — which people serving solely as technical advisers would hardly be able to require. Congress made clear that something really effective must be done about the torrent of Federal reports, and Circular No. A-40 reflects this intent.

Circular No. A-40 provides that a form which is approved by the Division of Statistical Standards shall include, on its face, approval language of this sort:

Budget Bureau No. 99-R999
Approval Expires January 31, 1950

The way in which clearance is shown for a regulation which lays down reporting or record-keeping requirements can be illustrated thus: If clearance has been given to a regulation which directs specified types of business firms to submit specified data semi-annually, the regulation includes a sentence which reads something like—"The reporting requirements of this regulation have been approved by the Bureau of the Budget in accordance with the Federal Reports Acts of 1942." The Bureau does not approve or disapprove any other parts of the regulation, such as any substantive aspects of export control, or a requirement that specified safety devices be used by railroads.

f. Experience under the Federal Reports Act.

The approval or disapproval of forms, etc., by the Budget Bureau, as required by Sec. 5 of the Reports Act, is a function which, above all others in that statute, needs to be delegated, as the actions under Sec. 5 are so numerous that final consideration of each separate one at the level of the Director would be impossible. In 1943, the first year after the enactment of the law, there were about 7,500 actions, of which 12% were disapprovals.

The disapproval rate is by no means the sole measure of success under the Federal Reports Act. The typical situation is neither approval "as is" nor complete disapproval, but approval after the D. S. S. and the submitting agency have jointly arrived at a reduction in the length or complexity of the form, a re-

duction in the number of people who will be asked to fill it out, an increase in the representativeness of the coverage, or some other improvement.

The disapproval rate has declined as agencies have improved their own standards. They prefer not to make proposals which will be rejected. This factor, plus the ending of most of the war-time economic controls, has brought about a marked decline in the aggregate number of forms, etc., submitted to the D. S. S. under Circular No. A-40.

What phases are coordinated?

On the whole, the statistical coordinating agency has concerned itself more with the plans for *collecting* data than with the plans for tabulating or analyzing them. Indeed, the accent in the Federal Reports Act is on the collection phase of statistical work. The reason is that Congress, in enacting this law, was motivated chiefly by a desire to have the public burdened as little as possible with filling out forms. The emphasis of the Division of Statistical Standards upon the collection phase was especially great during World War II, when the vast number of reporting devices resulting from economic controls made it impossible to devote much of the Division's effort directly to statistical phases other than collection.

Since the war, the Division has moved back toward a more comprehensive approach to statistical coordination. A reference has been made, above, to that part of the Congressional policy laid down in the Federal Reports Act which declares "that information collected and tabulated by any Federal agency should insofar as is expedient be tabulated in a manner to maximize the usefulness of the information to other Federal agencies and the public." The Division of Statistical Standards promotes the Act's policy, and especially this part of it, by taking an interest, formally and informally, in such matters as the shaping of tabulations. The interest may show itself in approving a data-collection proposal with the understanding that a specified scheme of tabulation is contemplated. In negotiating about this, as in negotiating with regard to the collecting phase, the Division of Statistical Standards is supported not only by the Federal Reports Act but also by its collaboration with the Division of Estimates (see above). Adequacy without waste is as essential in the tabulation and analysis phases as in *collecting* the data. As for analysis, however, the D. S. S. has held the view that here a competition of ideas is sometimes more useful than the avoidance of duplication.

The successive phases of any statistical venture are sufficiently interrelated that, to some degree, consideration of one necessarily involves considering the others. A decision as to whether it is justifiable to collect certain data cannot be sensibly made, nor does the D. S. S. attempt it, without attention to the questions, "What kind of tabulations, if any, are contemplated?" and "What sort of analysis, if any, is proposed?" The Division's integrating function in connection with the processing of data may either relate to a particular tabulation project, or be a broadly focussed undertaking such as the Standard Industrial and Commodity Classifications, developed under the Division's leadership. Tabulations based on these classification schemes are of use not only to the tabulating agency but also to other agencies which utilize these systems.

How the job gets done

The Division of Statistical Standards has never had a large staff, the total number of full-time members at present being fifty-odd, including thirty-odd professional people. This situation, as well as the Division's own predilections, has resulted in various procedures which are hardly discernible by studying the Federal Reports Act but which nevertheless promote the law's purposes. Wherever possible, the Division encourages an agency to establish within it a central clearance unit, having an intradepartmental role somewhat comparable to the Government-wide role of the D. S. S. Such a unit brings about much of the technical improvement which otherwise would be left to the D. S. S.; it causes some ill advised reporting schemes to die aborning; and it usually is sufficiently aware of the value of coordination that it helps to insure that, neither through inadvertence nor otherwise, does the agency do any statistical bootlegging—i.e., use a report form without D. S. S. approval. The D. S. S. has received invaluable assistance from its Advisory Council on Federal Reports (representing major business organizations), its Labor Advisory Committee on Statistics (representing the unions), and numerous interdepartmental committees, such as that on Standard Commodity Classification. When a proposal is made regarding the collection or processing or use of data, consultation with the responsible agency and with other interested agencies may be anything from a telephone call to a series of meetings of an *ad hoc* committee, depending upon the nature and importance of the problems. The integration of Federal statistics is a piece of teamwork in which the methods reflect the vast diversity of Government activities.

The Supplementary-Scale Chart

TWO CHARTS FOR THE PRICE OF ONE

by KENNETH W. HAEMER

In presenting time series data, there is often a question of whether to show the basic series or a significant ratio derived from it. For example: is it better to show dollar earnings, or earnings related to capital; supervisory wages, or supervisory related to total payroll; total radios, or radios per 100 families? Many times, the answer points very clearly one way or the other: but sometimes there is good reason to want to show both sides of the story, yet neither the space nor the emphasis assigned to the subject warrants the use of two charts. The supplementary-scale chart is one answer to this problem.

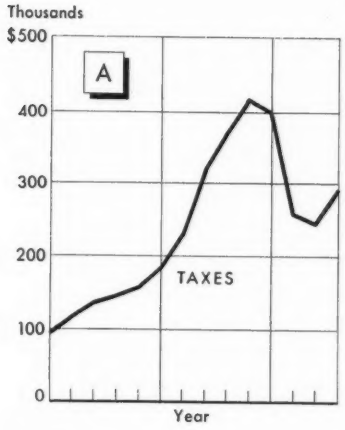
Charts A and B show two familiar absolute series, and Chart C the relationship of one to the other. A and C provide typical prospects for the combination type of picture illustrated by the supplementary-scale Chart D.

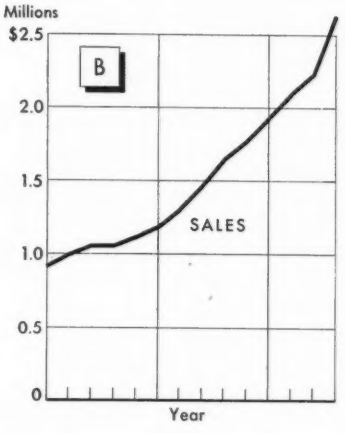
The procedure is extremely simple. Starting with Chart A, the first step is to determine the supplementary-scale range needed to include the highest and lowest points. From an inspection of Chart C, this range appears to be about 10 to 25%. The second step is to compute these percentages of curve B (plus as many in between percentages as are needed for ease in reading the chart), then, plot the resulting absolute series as a screen of light background curves.

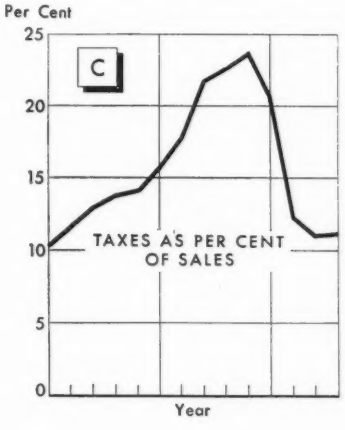
Series A remains in its original form, but is now supplemented by a flexible scale that shows the relationship at each point to series B, thus providing the gist of the information carried in series C. In fact, Chart D is almost the equivalent of three charts: although it doesn't directly show the magnitude of series B, it does reflect the general movement.

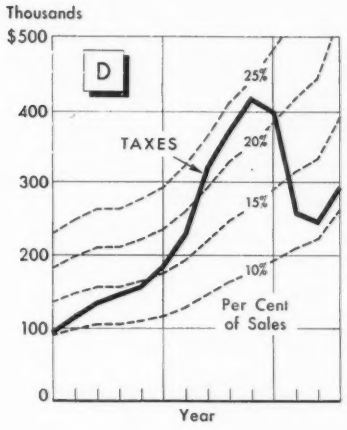
With some subjects, an added value can be squeezed from this type of chart. For example, suppose that Chart D showed "dollar earnings" against a supplementary scale representing 4, 6, 8 and 10 "percent of investment." It would then be possible to tell not only how much and what rate of return was earned each year; but conversely, how much should have been earned to yield any given rate of return.

Another interesting feature of this chart is its reversibility. Sometimes it can be worked either way—the basic series can be either A or B. For example, the curves might be "interest charges," and the background screen "percent of available income": or, the curve could be "income" and the supplementary lines "number of times interest earned." Inventories and sales, billing errors and bills submitted, automobiles and population, passenger miles and fatalities, are other typical pairs that can be switched in this way.









Two points of design should be mentioned: first, the supplementary scale lines should be very much subordinated; second, the labels for these lines should be set away from the major curve label and arranged one over the other in a compact column. Unless these procedures are followed, the chart is likely to be more confusing than enlightening.

Like all compromises, this two-in-one chart loses something: it does not pretend to do exactly as much work as the charts it replaces. Also, it has the disadvantage of a somewhat formidable appearance, a weakness that becomes more pronounced with every additional supplementary scale line. It certainly is not a chart for the general reader: in this as in most cases, the general reader will get the story better from two or three simpler charts than from one less simple one.

On the other hand, the supplementary scale adds important information that brings the basic series into proper perspective and points up the meaning or significance of the changes in this series. The combination presents a well rounded story in a neat and compact package, and is especially valuable as a space saver when the story must be told for a number of units such as a series of states, cities, companies, departments or plants.

In spite of its drawbacks, this chart does meet a specific limited need in an ingenious manner; and when properly conceived, designed, and employed, it is a useful form of presentation.

QUESTIONS and ANSWERS

edited by FREDERICK MOSTELLER

Solutions to Proposed Question A

Number of Different Kinds of Elements in a Population

Some time ago (AMERICAN STATISTICIAN, Vol. 3, No. 1, p. 23, 1949) we published a proposed question requesting solutions or references relating to it. The response from the readers has been so gratifying that we shall continue to propose questions to our readers when the problem seems particularly difficult or interesting. The problem, proposed by Charles Callard, was:

Proposed Question A. "A company has offered to give and has received requests for a free sample of its product. A large number of requests (say 100,000) have come in. These were filled. It is known that the same people often sent more than one request. From a sample of the requests how can we estimate how many *different* people have sent requests?

"More generally, we have k classes of similar objects, with n_1 objects in class 1, n_2 in class 2, . . . , n_k in class k. The class of an object is readily identifiable when the object is examined, but the classes are not ordered. We know $N = \Sigma n_i$. From a sample of size n drawn from the N objects, we wish to estimate k, the number of classes on the basis of the information given in the sample. We might view this statement as a problem of balls in an urn. In this case it is necessary to take a sample, and on the basis of population size and the sample which gives the numbers of say white balls, red balls, beige balls, etc., we are required to estimate how many different colors are present in the population."

At the time of the proposal the possibility of making a field study to find out the number of requests sent by single individuals was discussed. Also we discussed the shortcomings of the estimate Nk_s/n for the urn problem, where k_s is the number of classes observed in the sample.

William Haenszel, Director, Bureau of Vital Statistics, Department of Health of the State of Connecticut, Hartford, addresses himself directly to the problem of the requests for free samples. He suggests we draw a sample from the files and do a field study on the families represented by the individuals in the sample. He says: "The question on the number of different kinds of elements in a population has been discussed in a slightly different guise by Greenwood and Yule in an article entitled *On the Determination of Size of Family and of the Distribution of Characters in Order of Birth from Samples taken through Members of the Sibships*, published in the JOURNAL OF THE ROYAL STATISTICAL SOCIETY, 87, 2; 179 (January) 1914.

"One treatment of the problem propounded on the number of different people (families) requesting free samples would be as follows:

1) Follow-up investigation would determine whether a particular request drawn was from a family submitting a total of 1, 2, . . . , n requests.
2) A frequency distribution (F_x) of number of requests submitted by a family can be determined from the sample observed. The distribution is biased by the greater probability of inclusion in the sample of a family submitting many requests than of a family submitting one. To eliminate this bias, divided F_x by x.
3) The mean number of requests per family unit would then be

$$\bar{x} = \frac{\Sigma F_x}{\Sigma F_x/x}$$

The sample mean could then be divided into the total number of requests to estimate the number of family units."

Haenszel supplies the following illustrative example.

Example.

Requests Submitted	Distribution of Sample Observed (F_x)	Corrected (F_x/x)
1	800	800
2	100	50
3	60	20
4	40	10
Total	1000	880

$$\bar{x} = \frac{1000}{880} = 1.136$$

It should be noted that this estimate will give reasonable results — somewhere between the observed number of different people in the sample and the number of requests N. However, the method is not unbiased.

From the National Institute of Health, U. S. Public Health Service, Federal Security Agency, Jerome Cornfield writes that he, Samuel Greenhouse, Nathan Mantel, and Marvin Schneiderman teamed up to obtain the following results. Their contribution to the urn problem is that one can estimate the number of classes k from the number of different colors x observed in a sample of n drawn with replacement by solving

$$x = k - k\binom{N - N/k}{n} \Big/ \binom{N}{n}$$

for k in the special case where there are known to be the same number of balls of every color. This method gives reasonable but biased estimates. When sampling with replacement use solution of

$$x = k(1 - (1 - 1/k)^n),$$

and note that we need not know N.

For the practical problem where we draw a sample of n cards with replacement and follow up the x different individuals to discover how many requests N_i each sent, they suggest three possible estimates of k. First

$$x' = \sum_{i=1}^{k} a_i/[1 - (1 - P_i)^n],$$

where x′ is the estimate of k, a_i is 1 or 0 depending on whether the ith individual is in the sample or not. The $P_i = N_i/N$ are known for the individuals in the sample because we know N from our requests and the N_i from the followup. This estimate is unbiased, but the estimates may not be reasonable in the sense described earlier. Secondly they suggest

$$x'' = (N/n) \sum_{j=1}^{N} \alpha_j/N_j,$$

where α_j is 1 or 0 depending on whether the request is in the sample or not, and N_j is the total number of cards sent by the individual who sent the jth request. Note that this is not the naive estimate discussed in the original presentation of the problem. This estimate is also unbiased but can give unreasonable results. Thirdly they suggest a ratio estimate which is not unbiased but which they conjecture is preferable to either of the other estimates. The third estimate is

$$x''' = \frac{N \sum_{i=1}^{k} a_i/[1 - (1 - P_i)^n]}{\sum_{i=1}^{k} a_i N_i/[1 - (1 - P_i)^n]}$$

This estimate gives reasonable but biased results.

They point out the shortcoming that: "All three estimates yield an unsatisfactory result in the following situation: The sample of n requests yields X individuals, each of whom reports N_i and $\Sigma\, a_i N_i$ turns out to be exactly equal to N. In that case X is obviously the correct answer. None of the three estimates reduces to X in this case, however."

They also suggest a generalization: "As before we have k classes in the population, with N_i in the ith class. Each ball has a different number on it, the jth ball in the ith class having x_{ij}. The ith class is characterized by the number $\sum_{j=1}^{N_i} x_{ij}$ We now draw a sample of n balls and from it wish to estimate the k numbers, $\Sigma_j\, x_{ij}$, i = 1, 2,...,k. This is a formal statement of the problem of estimating the distribution of dividend recipients by size of dividend receipts from a sample of corporations, each corporation providing a statement of its dividend payments to its stockholders." The symbolic relations are:

k = number of different dividend recipients
N_i = number of dividends received by *ith* dividend recipient
x_{ij} = *jth* dividend check of *ith* person
$\sum_{j=1}^{N_i} x_{ij}$ = total receipts of *ith* person.

Leo Goodman, Department of Mathematics, Princeton Uni-

versity, has made an extensive investigation of the urn problem and has submitted a manuscript to the ANNALS OF MATHEMATICAL STATISTICS. Here we will outline the main results. If we draw without replacement a sample of n from a population of N balls containing N_i in the ith of k classes, then we can use the sample x_i to obtain the unique unbiased estimate of k, where x_i is the number of classes containing exactly i elements in the sample. This estimate may be written as

$$\hat{k} = \sum_{i=1}^{n} A_i x_i$$

where

$$A_i = 1 - (-1)^i \frac{[N-n+i-1]^{(i)}}{n^{(i)}}$$

and as usual

$$t^{(i)} = t(t-1)\ldots(t-i+1).$$

The only restriction is that the sample size must be larger than the largest number of elements in any one class. Furthermore Goodman has obtained unbiased estimates of k_j, the number of classes containing exactly j elements. These will not be presented here.

Although the estimate $\hat{k}$ is unbiased, it may give unreasonable results in the sense previously described. Goodman suggests alternative estimates which are more reasonable for the kinds of populations ordinarily presented in such problems. One of these is

$$\hat{k}' = \begin{cases} N - \dfrac{N(N-1)}{n(n-1)} x_2 = t & \text{if } t \geq \Sigma x_i \\ \Sigma x_i & \text{if } t < \Sigma x_i \end{cases}$$

For the problem using the field interview to find out the number of requests submitted by any individual in the sample, Goodman suggests the unbiased estimate

$$\hat{k}^* = \sum_{i=1}^{n} \frac{z_i}{1 - \dfrac{[N-n]^{(i)}}{N^{(i)}}}$$

or the approximation

$$\hat{k}^{**} = \sum_{i=1}^{n} \frac{z_i}{1-(1-p)^i}$$

where $p = n/N$, and z_i is the number of classes in the sample having i elements in the population.

To illustrate these various methods, we drew one sample of $n = 500$ both with and without replacement from each of two populations.

Population 1		Population 2	
k = 1000 classes of 10 each	Size of Class	Number	Frequency
	(1)	5000	5000
N = 10,000	(2)	2000	4000
	(3)	200	600
	(4)	100	400
		k = 7,300	10,000 = N

The samples drawn with and without replacement are nearly identical, 490 common elements for Population 1, 483 identical elements for Population 2. In the table initials at the left are those of the authors mentioned above.

ESTIMATES OF k BY VARIOUS METHODS

	Population 1		Population 2	
	Without Replacement	With Replacement	Without Replacement	With Replacement
No field Study	k = 1000	k = 1000	k = 7300	k = 7300
G. $\hat{k}'$	−128,000	−123,000 [1]	5992	−158,000 [1]
G. $\hat{k}$	399	392 [1]	6383	4213 [1]
C. et al	973	984	5537 [1]	4902 [1]
Field Study				
H. $N/\bar{x}$	1000	1000	7294	7303
C. et al x′	1014	996	7392	7171
C. et al x″	798	784	7148	6935
C. et al x‴	1000	1000	7256	7263
G. $\hat{k}^*$, $\hat{k}^{**}$	994	977	7216	7001

[1] Formula not appropriate.

Although two samples can scarcely be called definitive, we are pleased to find that in the case of the field study Haenszel's formula (the simplest of all) turns out best. When we work directly from the urn in Population 1, Cornfield's formula works best. Of course it has assumed the correct but ordinarily unknown fact that there are equal numbers in each class. However, it does not really require us to know N, which will be a help in some problems. Goodman's estimate $\hat{k}'$ is better for Population 2, but its behavior is extremely erratic, as witness its change of 2200 when applied to sampling with replacement from the same population. The unbiased estimate $\hat{k}$ is too erratic for any good.

To sum up: the Haenszel formula seems best for field studies, but none of these formulas seem really satisfactory over the whole range of populations for the urn problem. More work is clearly needed on the problem.

We have also received the following communication from C. F. Carter, Emmanuel College, Cambridge, England: "You may like to know that the same problem arose in finding the number of (different) investors in British industrial stocks. Here one can obtain a stratified sample from the share registers of companies, but it is a sample, not of investors, but of shareholdings. That is to say, a man who holds 100 (different) stocks is represented 100 times in the population from which the sample is drawn. In order to find the number of separate investors, I found the overlap between the sample and certain 'check registers' (i.e. the number of names in the sample of *all* shareholdings also appearing on the register of the chosen company) of large companies; and I made assumptions about (a) numbers *proportional to* the probabilities that an n-company investor ($n = 1, 2, \ldots, 100$) will invest in the check-register company, and (b) the proportion of investors who hold n securities. I proceeded by a method of Extreme Assumptions, making varying assumptions about the two quantities mentioned, and thus obtaining extreme limits within which the truth must lie.

"Thus, if T be the number of investors, q_n the proportion who own n stocks or shares, nTq_n is the number of shareholdings from which the sample is drawn. Let S be the number in the sample, C the number on the register of the large company used as a check, O the overlap between the two. Let p_n be the probability that an n-company investor will invest in the large company. Then, ignoring some minor corrections,

$$(1) \qquad T = \frac{C}{O} S \frac{\Sigma n p_n q_n}{\Sigma n q_n \Sigma p_n q_n}$$

and this is preferable to the simple form

$$(2) \qquad T = \frac{C}{\Sigma p_n q_n}$$

because it does not require us to know the absolute magnitudes of p_n.

"A rather inexact and journalistic account of our results will be found in the London FINANCIAL TIMES, March 2, 1949. You will see that the problem is similar to that you propound, but that it is soluble because the investors can be classified according to the companies they invest in."

It may not be quite obvious that the complicated estimate (1) given of the number of investors is a sensible one. It is clear, as the author states, that if the p_n employed in formula (1) are only proportional to the correct values instead of being the correct values, this estimate of T is unchanged. Formula (2) is obviously a sensible estimate of the number of separate investors because it merely divides the number of investors in the single company by the proportion of all investors who invest in that company. Consequently the right side of formula (1) is composed of a sensible estimate of T, namely the right side of formula (2), multiplied by a ratio which might be written

$$(3) \qquad R = \frac{S}{\Sigma n S q_n} \Big/ \frac{O}{\Sigma n S p_n q_n}$$

Here we have divided numerator and denominator by S merely for convenience of exposition. This ratio can be rewritten as

$$\frac{\text{persons per shareholding in sample}}{\text{persons per shareholding in register}}$$

where in the numerator the persons referred to are the S people in the sample, while in the denominator we are dealing with the subset of sampled people who are in the register. The two quantities are expected to be about the same so R should be near unity. This argument may help in understanding the rationale for formula (1).

Michigan University,
General Library,
Ann Arbor, Michigan.
A/S
PMF

CHAPTER NOTES

BOSTON

Officers for 1949-1950 are as follows: President—Roswell F. Phelps, Director, Phelps Statistical Service; Vice-President—William W. K. Freeman, Mutual Boiler Insurance Co.; Treasurer—E. L. Quirin, Babson's Statistical Organization; Secretary—Wilfred S. Lake, Northeastern University; Counselors—Frederick C. Mosteller, Harvard University and Louis C. Young, Massachusetts Institute of Technology.

CHICAGO

Warren N. Cordell, Vice President and Oliver S. Castle of A. C. Nielsen Company summarized new ideas developed from their most recent experiences in applying probability sampling methods to the problems arising in market research at the luncheon meeting on May 12.

At a dinner meeting on May 17, Mr. Henry B. Arthur, Economist, Swift & Company and former Chief of the Office of Program Review of the E. C. A., in Paris, discussed "Economic Aspects of the Marshall Plan". Mr. Arthur's presentation was concerned with the objectives, progress and problems of the European Recovery Program.

At a concluding meeting of the Statistical Techniques Luncheon Group, Professor W. Allen Wallis, University of Chicago, described an estimation of population values by sequential methods.

Maxwell R. Conklin, at the May 26 meeting of the Chapter, told of the types of published and unpublished data to be made available on the 1947 Census. The results of the Census, currently being tabulated, show the vast changes since 1939 and provide new benchmarks for business planning.

In accordance with unbroken precedent, the last Chapter meeting of the season on June 22 was devoted to "The Business Outlook". Members of the Chapter were given the opportunity to make their own predictions. Walter Hoadley gave a summary and analysis of the Chapter's forecasts for the previous year. Paul McCracken, Associate Professor of Business Conditions, University of Michigan and Samuel Weiss, Chief of Employment Statistics, U. S. Bureau of Labor Statistics, were guest speakers for the evening.

The following officers were elected by the Chicago Chapter for the year 1949-1950: President—Robert J. Eggert, American Meat Institute; Vice-President (in charge of evening meetings)—William W. Tongue, Jewel Tea Company, Inc.; Vice-President (in charge of luncheons)—Lucile Derrick, University of [illegible] of membership)—Leo J. O'Grady, Illinois Bell Telephone Co.; Secretary—Charles G. Wright, Federal Reserve Bank of Chicago; Treasurer—Marian J. Richeson, The Peoples Gas Light and Coke Company; Directors—Willard D. Arant, Swift & Company; Adolph O. Berger, U. S. Bureau of Labor Statistics; Lester S. Kellogg, Deere & Co.; George A. Morgan, The Peoples Gas Light and Coke Co.; Robert H. Perry, U. S. Department of Commerce; De Ver Sholes, Chicago Association of Commerce & Industry; W. Allen Wallis, University of Chicago; Mason Wescott, Northwestern University; Chapter Representative on District Committee—Walter E. Hoadley, Jr., Federal Reserve Bank of Chicago.

CENTRAL NEW JERSEY

"The Use of Statistical Information in Research and Administration" was the subject of the May 26 meeting of the Chapter. Frederick F. Stephan, Director of the Study of Education at Princeton and Professor of Social Statistics, viewed the work of the statistician with special attention to major problems of producing and utilizing statistical information.

CENTRAL INDIANA

The Central Indiana Chapter concluded its winter program with the election of officers for the year 1949-1950: President—Irving Burr, Purdue University; First Vice-President—George B. Haines, Indiana University; Second Vice-President—Roy Adams, Department of Public Welfare; Secretary-Treasurer—Lucy Kantz, Indiana Employment Security Division.

Mr. John Hanson, retiring President, was speaker for the evening. He dealt with "Techniques in Sales Forecasting".

COLUMBUS

The following officers were elected by the Chapter for the term 1949-1950: President—David W. Lattimer; First Vice-President—Ivon W. Ulrey; Second Vice-President—Wharton F. Keppler; Secretary-Treasurer—Mikhail Condoide.

CONNECTICUT

A dinner meeting was held on May 13 at which the speaker was John D. Norton of the Planning Research Division, U. S. Air Force. Mr. Norton spoke on "The Statistical Basis of Military Management".

CUBAN STATISTICAL SOCIETY

A meeting to reorganize and reactivate the Cuban chapter of the ASA was held in Havana on May 26. The following officers were elected pending the revision of the Chapter constitution and subsequent elections: President—Ing. Hugo Vivo; Vice-President—P. A. Francisco Sala Perera; Treasurer—P. A. Jaime Proaza; Secretary—Dr. Ulises Giberga.

A committee to revise the chapter constitution in conformity with the constitution of the ASA was organized, consisting of Hugo Vivo, Ramon Gutierrez Gammoneda, and Ulises Giberga.

DENVER

The final meeting of the season, held on May 12, was addressed by Dickson H. Leavens, former managing editor of Econometrica on the subject of "Silver in Two Wars".

PHILADELPHIA

Benedict Saurino addressed the Philadelphia Chapter at a dinner meeting on May 20. Mr. Saurino spoke on "A Petroleum Industry Analysis".

OKLAHOMA CITY

Dewitt Hunt, Head of the Department of Industrial Arts Education and Engineering Shopwork at Oklahoma A. and M. College spoke before the May 4 meeting of the Chapter. Dr. Hunt discussed the subject of "Accidents" and supplemented his talk with data sheets giving comparisons of accidents in Oklahoma for the last five years.

ST. LOUIS

"Plans for the 1950 U. S. Census" were discussed at the June 7 luncheon meeting of the St. Louis Chapter. Howard G. Brunsman, Chief, Population Division, Bureau of the Census, and Morris Hansen, Assistant Director for Statistical Standards, Bureau of the Census, spoke on the types of data to be collected in the forthcoming 1950 census, areas for which data will be available and the tabulations which will be published regarding the material obtained.

UNIVERSITY OF ILLINOIS

The following officers were elected at the meeting of May 9 for 1949-1950: President—Margaret G. Reid; Vice-President—David R. Saunders; Secretary-Treasurer—Robert Ferber.

WASHINGTON STATISTICAL SOCIETY

The Chapter reports the election of the following officers for the period 1949-1950: President—Jerome Cornfield, National Institute of Health; Vice-President—Ezra Glaser, Bureau of the Budget; Secretary-Treasurer—Margaret Gurney, Bureau of the Census.